TASK 1

Describe the linear regression model. The answer should be comprehensive. The practical example with data and the
python code are mandatory. Place your answer into the task1.pdf file. The source code and the data should be included
in the resulting pdf file. No additional files are required.

ANSWER:

Linear regression is a fundamental statistical method used for modeling the relationship between a dependent variable and one or more independent variables. It assumes a linear relationship between the independent variables and the dependent variable.

1. Dependent Variable (Y): The variable we want to predict or explain.
2. Independent Variable(s) (X): The variable(s) used to predict the dependent variable.
3. Linear Relationship: The relationship between the independent and dependent variables is assumed to be linear, meaning the change in the dependent variable is proportional to the change in the independent variable(s).
4. Regression Line: The line that best fits the observed data points, representing the relationship between the independent and dependent variables. It is defined by the equation: $Y = Beta0+Beta1X+e$
   where Y is the dependent variable, X is the independent variable, Beta0 is the intercept , Beta 1 is the slope, and epsilone is the error term representing the difference between the observed and predicted value
5. Ordinary Least Squares (OLS): The most common method for estimating the parameters of the regression line by minimizing the sum of the squared differences between the observed and predicted values.

PRACTICAL EXAMPLE:


House Price vs. Size
Price ($)
300000
280000
260000
240000
220000
200000
1100
1200
1300
1400
1500
1600
1700
1800
1900
Size (sq. ft)
x=1556.2 y=3.157e+05